EO Clinic project:

Mitigation of Climate Change Risks in the Agricultural Sector of Cambodia

Work Order Report

Support requested by:
United Nations Development Programme (UNDP)

Reference: EOC0010_WOR_ v01
Date: 2020 September 21

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

REFERENCE DOCUMENTS

[RD-1]	ESA Request for Proposal: EOC0010_RFP_v02
[RD-2]	Technical Proposal: EOC0010_PRO_C_T_v01 by GeoVille Gmbh and SIRS
[RD-3]	Financial Proposal: EOC0010_PRO_C_F_v02 by GeoVille Gmbh and SIRS

ABOUT THIS DOCUMENT

This document is the final Work Order Report of the ESA EO Clinic project EOC0010 *Mitigation of Climate Change Risks in the Agricultural Sector of Cambodia*.

This publication was prepared in the framework of the EO Clinic (Earth Observation Clinic, see below), in partnership between ESA (European Space Agency), the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) / GIZ Regional Economic Development Program IV (RED IV) and a team of service providers contracted by ESA: GeoVille GmbH (Austria) and SIRS (France).

This Work Order Report (WOR) is structured as in the following:

- **Section 1** describes the context of GIZ's activities on Mitigation of Climate Change Risks in the Agricultural Sector of Cambodia, the project objectives and requested EO products and services.
- **Section 2** highlights the applied work logic and responsibilities among the EO Clinic service providers.
- **Section 3** describes the services and products provided, their specifications, methods and outcomes.

ABOUT THE EO CLINIC

The EO Clinic (Earth Observation Clinic) is an ESA (European Space Agency) initiative to create a rapid-response mechanism for small-scale and exploratory uses of satellite EO information in support of a wide range of International Development projects and activities. The EO Clinic consists of "on-call" technically pre-qualified teams of EO service suppliers and satellite remote sensing experts in ESA member states. These teams are ready to demonstrate the utility of satellite data for the development sector, using their wide range of geospatial data skills and experience with a large variety of satellite data types.

The support teams are ready to meet the short delivery timescales often required by the development sector, targeting a maximum of 3 months from request to solution.

The EO Clinic is also an opportunity to explore more innovative EO products related to developing or improving methodologies for deriving socio-economic and environmental parameters and indicators.

The EO Clinic was launched in March 2019 and is open to support requests by key development banks and agencies during the 2 years project duration.

AUTHORS

The present document was prepared and coordinated by Christian Treml (Framework Coordinator, GeoVille) with support from the following contributors: Björn Dulleck (Technical Lead Service 1, GeoVille), Norman Kiesslich (Head of International Development, GeoVille), Antoine Masse (Technical Lead Service 2, SIRS).

ACKNOWLEDGEMENTS

The following colleagues provided valuable inputs, insights and evaluation feedback on the work performed: Dr. Stefan Hanselmann (Program Director Regional Economic Development Program IV) and Zoltan Bartalis (ESA Coordinator and Technical Officer).

For further information	Please contact Christian Treml, Framework Coordinator, GeoVille (treml@geoville.com) with copy to Zoltan.Bartalis@esa.int if you have questions or comments with respect to content or if you wish to obtain permission for using the material in this report. Visit the ESA EO Clinic: https://eo4society.esa.int/eo_clinic.

1 DEVELOPMENT CONTEXT AND BACKGROUND

1.1. Climate Change Risks in the Agricultural Sector

Cambodia is highly vulnerable to the effects of climate change. It is facing a climate condition which presents increasing extreme weather events and negative impacts in the forms of casualties and obstructions to the country's economic growth and development, especially on those living in rural and remote areas where subsistence agriculture and natural resources are their main sources of livelihoods and local economy.

The Regional Economic Development (RED IV) program, co-financed by Switzerland and GIZ Germany, aims at strengthening the capacity of subnational and local governments and support rural poor to increase their income and actively participate in local economic development, thereby reducing poverty of the rural population. Hard-won rural improvements are fragile as climate change is an increasingly serious threat facing many communities in rural Cambodia. Main concern moving forward is rising temperatures, erratic weather patterns, and water shortages.

Thus, water management is an important issue for the RED IV program, which has been working on an analysis of the local availability of surface water in the north-western provinces of Cambodia since October 2019. Data availability in Cambodia is a challenge, as data are either out of date or not accurate enough.

Technical support is required by RED (IV) to improve flood and water resource management for the two provinces Oddar Meanchey and Kampong Thom. Within the framework of a comprehensive water resources management approach, information is required on the mapping of surface waterbodies as well as wetlands and their associated dynamics.

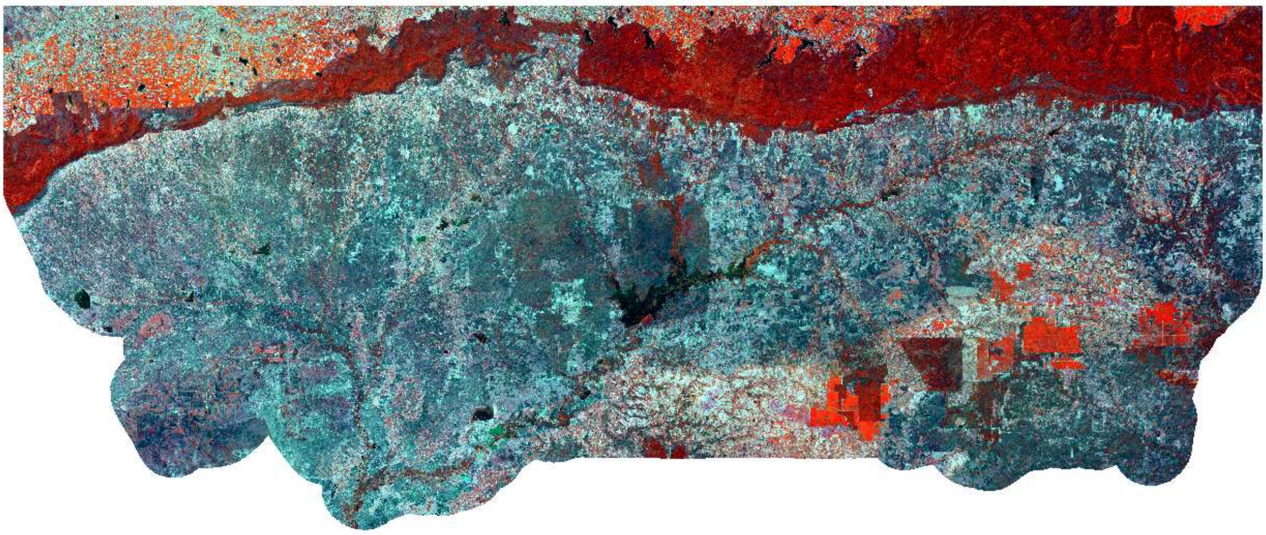


Figure 1: False colour Sentinel-2 image composite - Oddar Meanchey province (Feb. 2020) (Source: ESA)

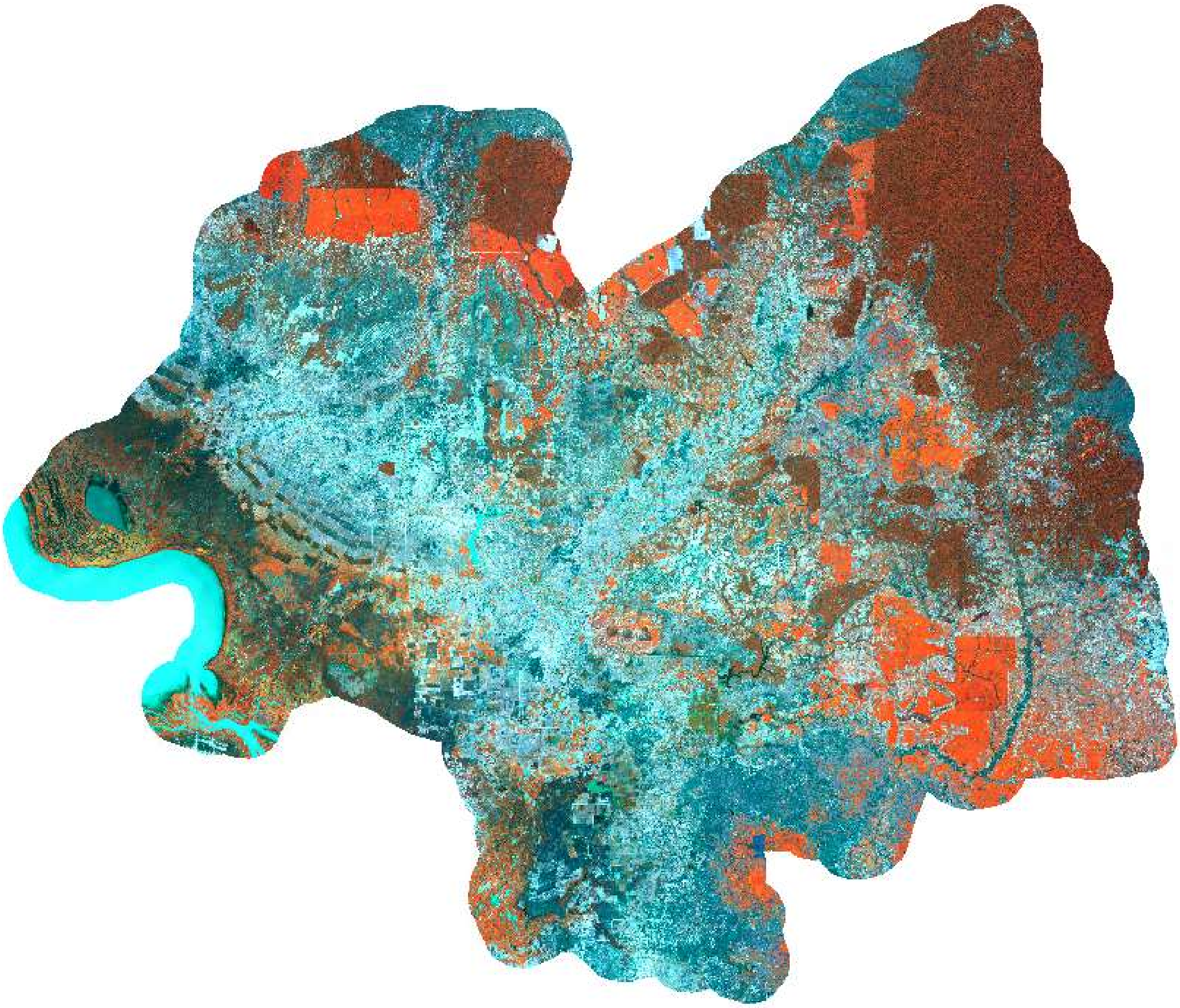


Figure 2: False colour Sentinel-2 image composite - Kampong Thom province (Feb. 2020) (Source: ESA)

1.2. Objectives

In response to these requirements and challenges, GIZ has requested Technical Assistance from the ESA EO Clinic. The need for proper and timely information on water (non-) availability is the most important requirement for water management activities. Hence, the aim of this EO Clinic activity is to support GIZ on the mapping and monitoring of surface waterbodies / wetlands and the assessment of their dynamics with a focus on the two provinces Oddar Meanchey and Kampong Thom.
The mapping shall demonstrate the capabilities and limitations of satellite Earth Observation (EO) data for an independent water / wetland monitoring and improve flood and water resource management for the two provinces.

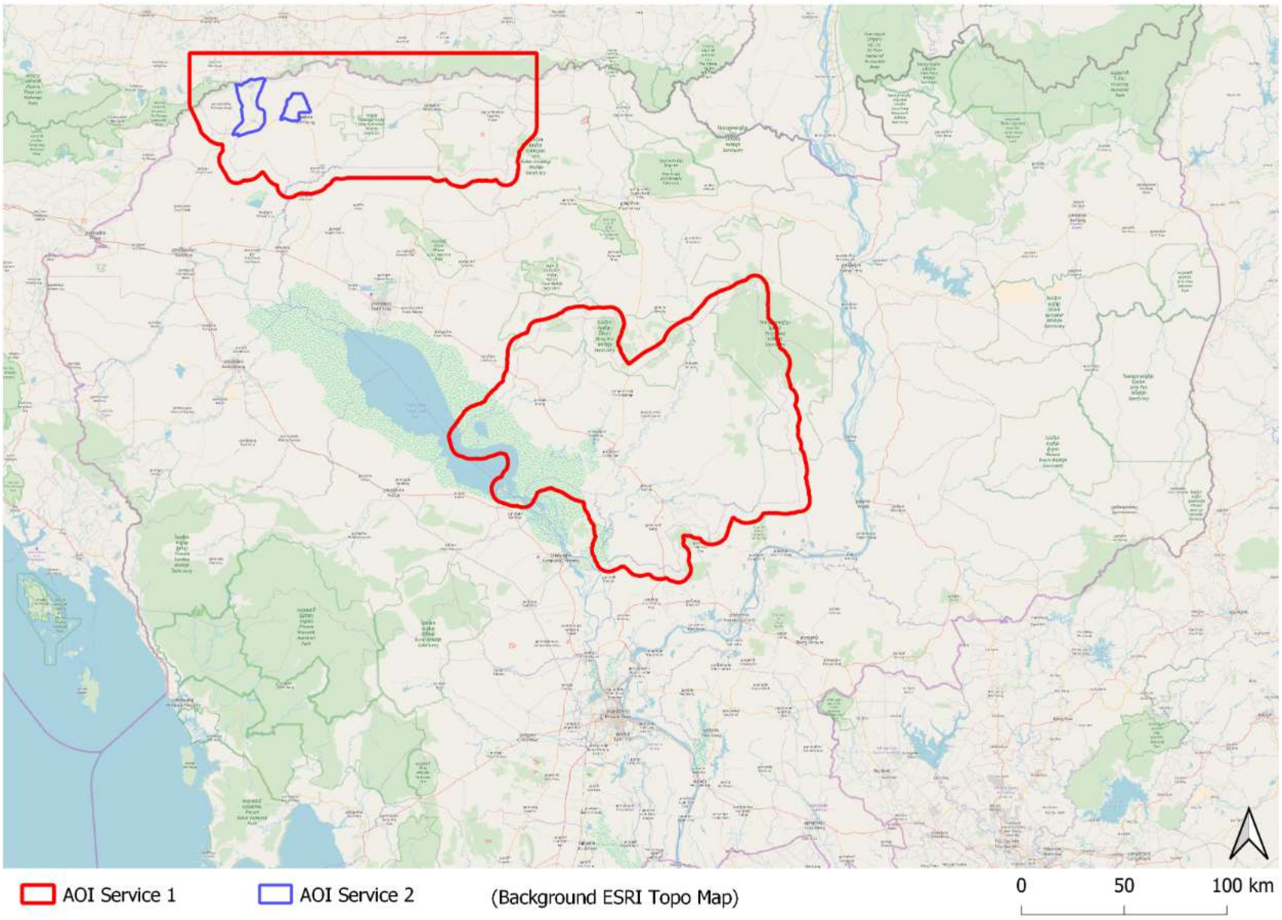


Figure 3: Service 1 and 2 - Areas of Interest

The following information services have been requested:

- **Service 1: Waterbody Inventory and Dynamics** – mapping of surface waterbodies (streams, rivers, lakes, reservoirs, creeks, irrigation canals) and wetlands as well as monitoring of the dynamics / temporal fluctuations in waterbody extent, taking into account the variability between wet and dry seasons.
- **Service 2: Detailed Waterbody Inventory** – mapping of small water features, such as narrow irrigation canals, in complement to the products of Service 1 based on very-high-resolution imagery (VHR) for two communes Kouk Mon and Bansay Reak within province Oddar Meanchey.

The initial set of information products and insights generated within this activity shall be used mainly by the GIS expert currently working with RED IV, the provincial and district administrations (Departments for Water Management, Departments of Agriculture). Moreover, outcomes shall serve to further encourage the adoption of EO within GIZ and other stakeholders.
The goal is to upscale the EO contributions and transfer the lessons learned to other GIZ programs in Cambodia working with agriculture in different provinces – depending on the interest of GIZ.

2 WORK LOGIC

The overall work logic and organisation between GeoVille (coordinator of EO-Clinic framework contract, interface towards ESA and GIZ, and service provider) and SIRS (service provider) is presented in the below. GeoVille has acted as focal point towards ESA as the contracting authority and GIZ Cambodia.

During the implementation phase the following stakeholder representatives were involved:

- Dr. Stefan Hanselmann, Program Director Regional Economic Development Program IV

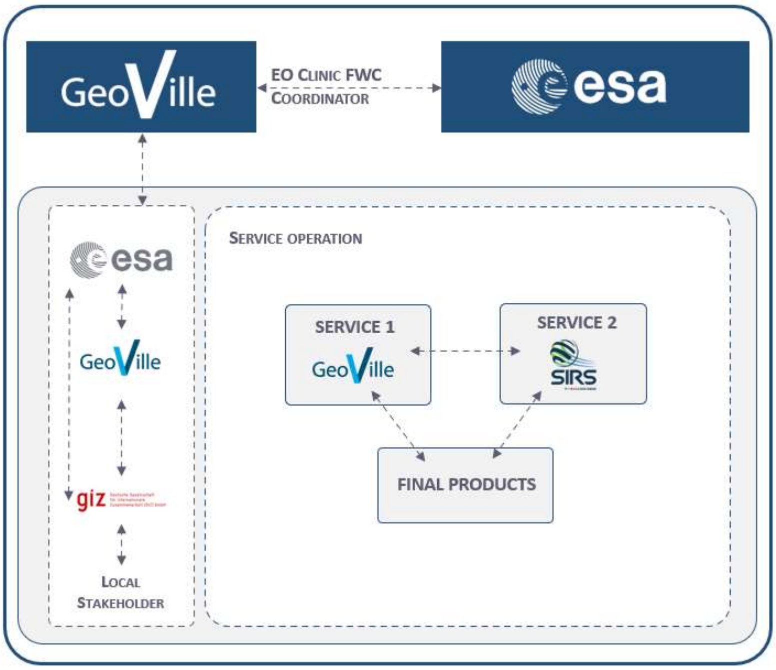


Figure 4: Overall Work Logic and interaction between organisations and services

Service 1 was implemented by GeoVille and Service 2 was generated by SIRS.

This work was initially planned over a tight period with a Work Completion Deadline (WCD) for both services of 5 weeks from the issuing of the Work Order (WO). Due to late kick-off meeting, vacation time and coordination of VHR data acquisition the Work Completion Deadline was extended in agreement with all stakeholders to September 21, 2020.

3 DELIVERED EO-BASED PRODUCTS AND SERVICES

3.1 Service 1 – Waterbody Inventory and Dynamics

3.1.1 Specifications

The technical specifications of the products adhere to the proposed properties in the technical proposal. This includes the following water products for each province (Oddar Meanchy and Kampong Thom): 39 Monthly Water Masks, 6 Water Frequency maps, 6 Minimum Water Extent maps and 6 Maximum Water Extent maps. In addition to the waterbody maps, 3 AnnualWetland Maps were delivered.

All products are based on Sentinel-1 and Sentinel-2 in 10m resolution.

Table 1: Service 1 product list

Product	Count	Pixel size	Reference period	File name example
Monthly Water Masks	39	10m	monthly	OM_waterMask_2019_05.tif KT_waterMask_2019_05.tif
Water Frequency	6	10m	seasonal	OM_waterFrequency_dryseason_18.tif * OM_waterFrequency_wetseason_18.tif
Minimum Water Extent	6	10m	seasonal	KT_minimumWater_dryseason_20.tif * KT_minimumWater_wetseason_20.tif
Maximum Water Extent	6	10m	seasonal	OM_maximumWater_dryseason_20.tif * OM_maximumWater_wetseason_20.tif
Annual Wetland Maps	3	10m	annual	KT_wetlandMap_2017

* seasons-definition:

- the Dry Season products cover the yearly period from October - April of the following year – product name refers to the end-year (2018) of the Dry Season.
- The Wet Season products cover the period from May – September within the same year

A detailed overview of the satellite-based input data, the data specifications and thematic information is provided in Table 2 below.

Table 2: Service 1 Product specifications – Waterbodies / Wetlands

General	
Resolution and Data Input	10m – Sentinel 1 and Sentinel 2
Geographic Projection	UTM Zone 48
Format	GeoTIFF
Datatype	Byte
Thematic information	
Classes and Codings	**<u>Monthly Water Masks</u>** - 001 Water - 000 No Water - 255 No Data

	Minimum Water Extent - 001 Minimum water extent - 000 No Water - 255 No Data **Maximum Water Extent** - 001 Maximum water extent - 000 No Water - 255 No Data **Water Frequency Maps** - 000% water to 100% water **Wetland Maps** - 000 Dry - 001 Permanent Water - 002 Temporary Water - 003 Permanent Wet - 004 Teporary Wet - 255 No Data
Accuracies	
Geometric positional accuracy:	Sub-pixel (<10m)
Overall thematic accuracy:	>90%
Minimum Mapping Unit (MMU)	
Minimum Mapping Unit (MMU)	2x2 pixels (400m²)
Minimum Mapping Width	1 pixel (10m)

Visual examples of the delivered products can be seen in Figures 5-8:

Figure 5: Monthly Water Mask

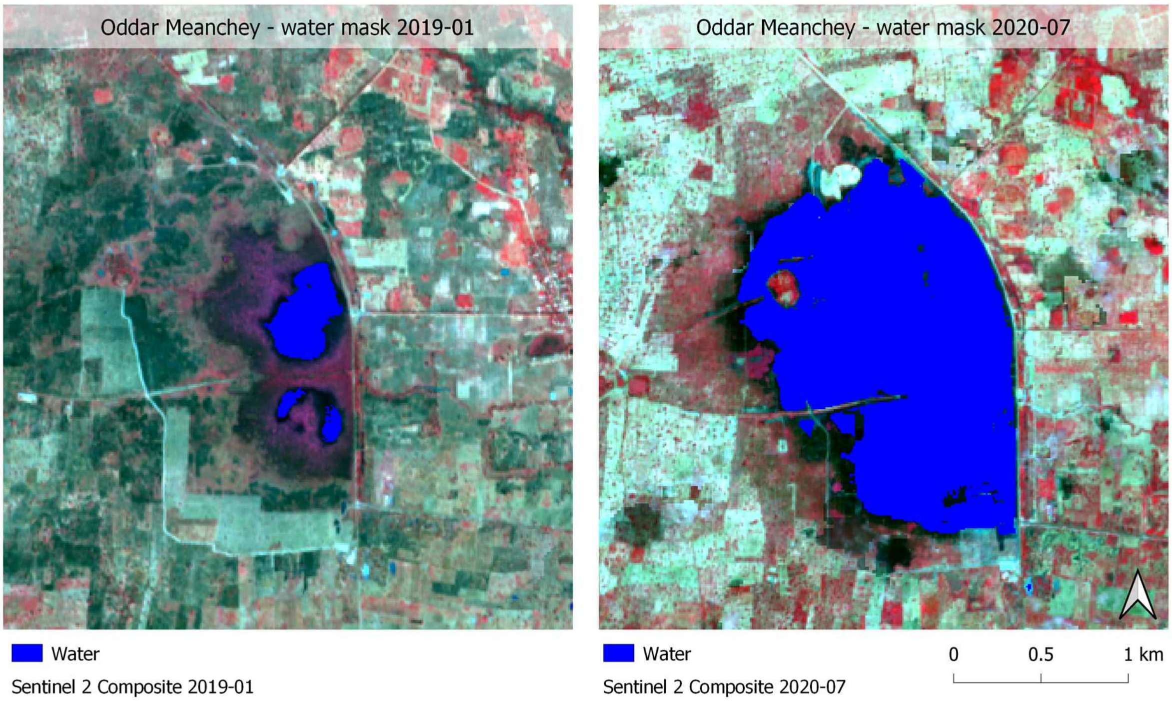


Figure 6: Water Frequency map

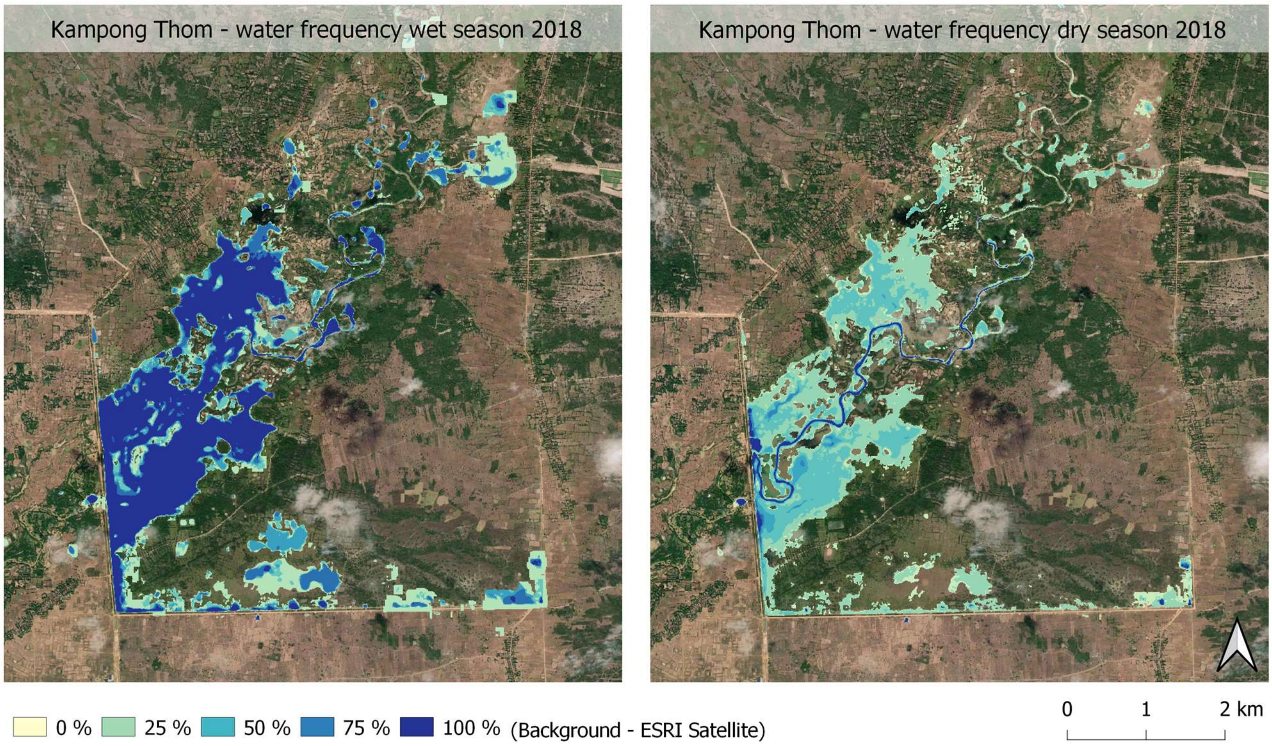


Figure 7: Minimum/Maximum Water Extent maps

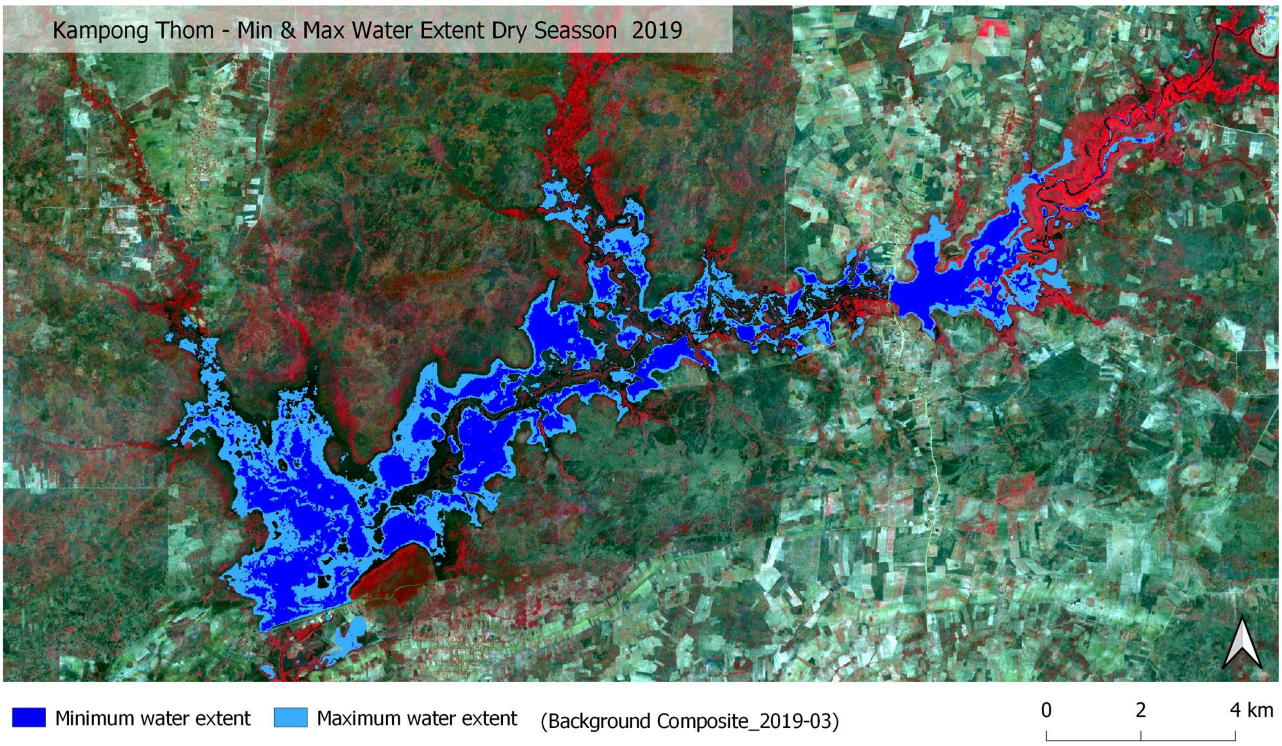
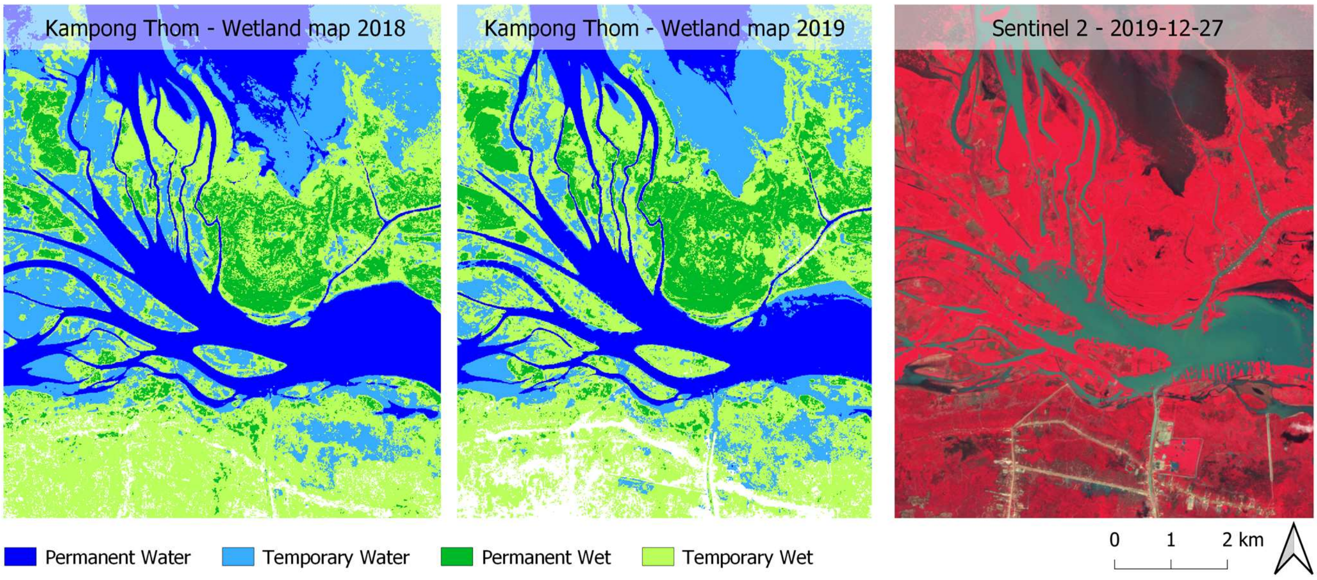


Figure 8: Annual Wetland map



Input Data

Generally, two input data types were distinguished for the water and wetland products. These consist of optical satellite imagery derived from the **Sentinel-2** Archive and **Sentinel-1** Radar imagery. In addition, **SRTM** DEM data were used to derive topographic indices.

Availability and limitations:

Sentinel-2B data is available from July 2017 onwards which can have an impact on the quality of the products from former dates. SCL masks supplied with Sentinel-2 L2A can have higher omission errors especially for cloud shadows and thin clouds.

Methods used

Pre-processing:

Sentinel-1 data were pre-processed using the ESA SNAP engine by applying:

- Orbit geometry corrections
- Border noise removal
- Radiometric calibration
- Speckle filtering
- Terrain correction.

ESA does not provide Sentinel-2 L2A imagery across Cambodia for the entire time series from 2017-2020. Data which has been ingested before that time is available as L1C only, thus, the correction of atmospheric influences (gaseous and molecular scattering and absorption, aerosol scattering and absorption) as well as cloud and cloud shadow masking had to be completed. The most recent Sen2Cor version (2.8) was used for atmospheric correction and cloud detection for data captured prior to May 2017 and December 2018, aiming at a maximum consistency among the entire Sentinel-2 time series.

Additionally, a multitemporal cloud shadow detection algorithm (Ludwig et al, 2019) was applied to reduce the Sen2Cor cloud shadow omission error rate.

For monthly compositing, the geometric median is used instead of a best-pixel approach (BPA). BPA composite methods only select the 'best' pixel of the whole time series, whereby valuable information can be lost (e.g. changes in water level or vegetation cover). Contrasting this, statistical based methods taking all valid observations into account are more representative than a single time step of the time series. Geometric median compositing is a suitable approach, as it is "using a high-dimensional summary statistic that applies to all bands at once to guarantee that the biophysical relationships among all spectral bands are maintained" (Roberts et al, 2017). In contrast to simple summary statistics such as the mean, the geometric median has a higher breakdown point (0.5) and is more robust against outliers, e.g. undetected cloud pixels. In addition to the more meaningful pixel values, the geometric median yields to smoother composites with less noise and artefacts.

Water Detection:

Based on the corrected and combined optical imagery, spectral indices and biophysical parameters such as Normalized Difference Water Index (NDWI) and Modified Normalized Difference Water Index (MNDWI) were derived from each monthly composite due to its sensitivity to surface water.

Sentinel-1 time series backscatter have been reduced to monthly metrics such as minimum backscatter to capture all monthly water occurrences.

Within the optical domain, water is detected based on the enhancement of the spectral signature using biophysical indices sensitive to water and including automated dynamic image thresholding methodologies (Martinis et al., 2009) as described in Ludwig et al. (2019).
The radar-based algorithm uses a pixel-based adaptive thresholding approach of the radar backscatter signals applying a dynamic variable backscatter threshold for each pixel based on statistical information from the adjacent pixels. The resulting water mask is then compared against the Height Above Nearest Drainage (HAND) index (Rennó et al., 2008) to exclude unrealistic water occurrences.
After processing the optical and radar imagery separately, the data were fused into monthly surface water extent masks – thus the advantages of both sensor systems could be used. To reduce commission errors, masking was applied using a potential water mask derived by percentile statistics on the aggregated water indices over the whole time series (2017-2020).

Wetness Detection:

Surface soil moisture sensitive biophysical spectral indices such as the Normalized Difference Moisture Index (NDMI) or the Angle Based Drought Index (ABDI) have been calculated, grouped and aggregated to monthly wetness probabilities sensitive to bare soil, sparse and dense vegetation moisture. As in the case of water detection, dynamically Otsu thresholding was applied to the probabilities to derive wetness masks for each month.

Classification:

Based on the cumulative water, wet and dry masks, and the total number of valid monthly observations, frequencies (in percent) are calculated for water and wet (soil, dense vegetation, sparse vegetation). These relative frequencies form the basis to swiftly generate the thematic classification of permanent and temporary water/wetness classes.

The products were derived from the intermediate cumulative water/wet/dry occurrence layers and their associated relative frequency layers (i.e. cumulative occurrence layers in relation to the total number of valid observations). The relative frequencies form the basis for the classifications according to the rules summarized in the tables below.

Wetland Map (annual):

Code	Class	Water relative frequency	Wet relative frequency
1	Permanent water • always water	> 80% Water	
2	Temporary water • alteration of dry and water • alteration of wet and water • with varying degrees of wetness • water instances dominate over wet	>25 - 80% Water	Water > Wet
3	Permanently wet areas • always wet		> 80% Wet
4	Temporary wet areas • alteration of dry and wet • with minor instances of water • wet instances dominate over water	< 80%	Wet > Water

Table 3: Classification criteria for the Wetland product

Minimum and maximum water extent (seasonal):

Code	Class	Water relative frequency
1	Minimum water extent • always water	> 80% Water
1	Maximum water extent • alteration of dry and water	> 10% Water

Table 4: Classification criteria for the Minimum and Maximum water extent products.

3.1.2 Quality Control and Validation

We applied a stratified random sampling approach to assess the accuracy of the water mask product. To extract the sample points in each stratum, the "LACOVAL – validation tool for land cover and land cover changes" has been used. This tool was developed within an ESA funded project in 2013 and is available open source. For the underlying validation the ArcGIS add-in of LACOVAL has been applied to the data. LACOVAL ensures that the sample selection and quality assessment follow a standardized protocol.

The first part of the validation consists of creating validation sample points, for which the samples were randomly selected (by the LACOVAL tool) in each stratum. After creating the stratified random samples, each point needed to be visually interpreted by an operator. This interpretation can be done in two ways, either *blind* (operator has no information about the classification and assigns a LC class to the reference point) or *plausible* (operator has to decide – yes or no – whether a reference point belongs to particular class without knowing the underlying classification). For the validation herein, we applied the *blind* approach.

The point by point interpretation is based on particular reference data sets and/or the inspection of the production data (if no appropriate reference data set is available). The choice of the reference data is depending on the resolution of the data being validated. According to the accuracy of the product either images with a low, medium, high or very high (VHR) resolution are used, also aerial photographs are possible. In addition, spectral indices can support the decision making in tricky situations, for which the land cover is not exactly identifiable from the HR/VHR imagery. After dealing with all sample points, LACOVAL is generating a validation report, yielding the producer's, user's and overall accuracy including the 95% confidence interval.

The number of validation points has been selected according to best practice guidelines (e.g., Congalton and Green, 2009; Olofsson et al., 2014). Validation was performed at the Oddar Meanchey site on the monthly water mask of January 2019.

Table 5: Summary of the accuracy assessment for Service 1

	Reference Data				
Classification	No water	Water	Row Total	User's Accuracy	95% Confidence Interval
No water	348	2	350	99,43%	0,40
Water	5	45	50	90,00%	4,67
Column Total	353	47	400		
Producer's Accuracy	98,58%	95,74%			
95% Confidence Interval	0,63	3,05			

Overall Accuracy	98,25%
Overall 95% Confidence Interval	0,65
Kappa	0,92

3.1.3 Outputs & Analyses

3.1.3.1 Outputs

The first category of outputs for Service 1 are the Monthly Water Masks for the provinces Oddar Meanchey and Kampong Thom each and for the time period May 2017 – July 2020. The second category of outputs is the Seasonal Product Suite including the following products: Water Frequency, Minimum Water Extent and Maximum Water Extent. The third category of outputs are the Annual Wetland Maps.

Delivery package (for each province):

- 39 Monthly Water Masks
- 6 Water Frequency maps
- 6 Minimum Water Extent maps
- 6 Maximum Water Extent maps
- 3 Annual Wetland Maps
- Metadata.xml files

3.1.3.2 Analyses

An initial analysis of the water time series (monthly masks) was performed to check the three-year series for significant trends in the lateral extent of water bodies in both provinces. Figure 9 shows the water extents of both regions. The curve for Kampong Thom (KT) clearly follows a seasonal pattern that appears less pronounced for Oddar Meanchey (OM). With an average of about 800 km2, water covered areas in KT are larger than those in OM by about an order of magnitude with around 90 km² for OM. This is mostly due to the inclusion of larger swaths of permanent water bodies to the south-west as a result of a buffer that was applied to the area of interest, but also, to a lesser extent, due to the regular flooding of large parts of agriculturally used lands, presumably for the cultivation of rice, during the growing season. Each of the areas are illustrated separately in Figure 10 and Figure 11 and the area values by month are listed in APPENDIX B.

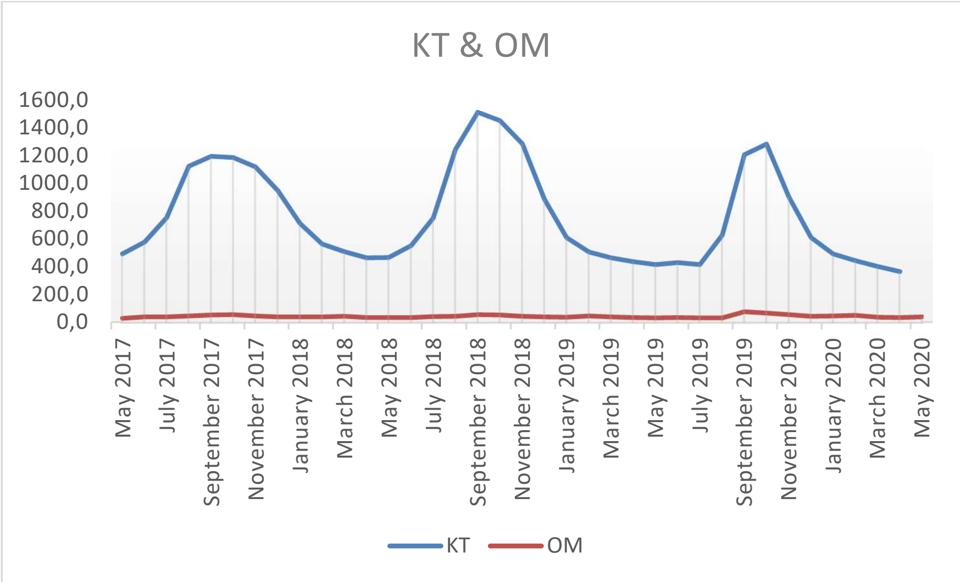


Figure 9: Water extents in km² for both provinces over the period of three years.

Fitting a linear trend through the 3-year series for province Oddar Meanchey reveals a significant negative trend, implying that water resources become scarcer at least during the observed period. The seasonal increase in surface water extents toward the end of the calendar years appear to become more pronounced with steeper increases and decreases before and after the maximum. A sample size of three years is low and extending the series back into the past as well as continuous monitoring in the future will help assess whether these findings are limited to the years 2017 to 2020 or part of a longer term trend.

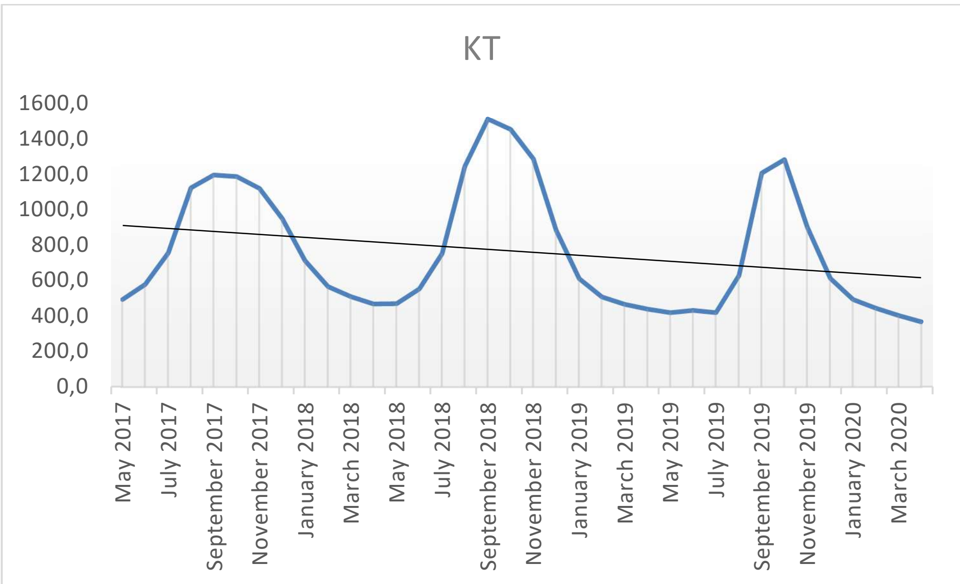


Figure 10: Water extents in km² for Kampong Thom over the period of three years with a linear trend.

The seasonal trends in the surface water extents in the Oddar Meanchey province are significantly less pronounced than for Kampong Thom. There is a slightly positive trend though that may not be statistically significant.

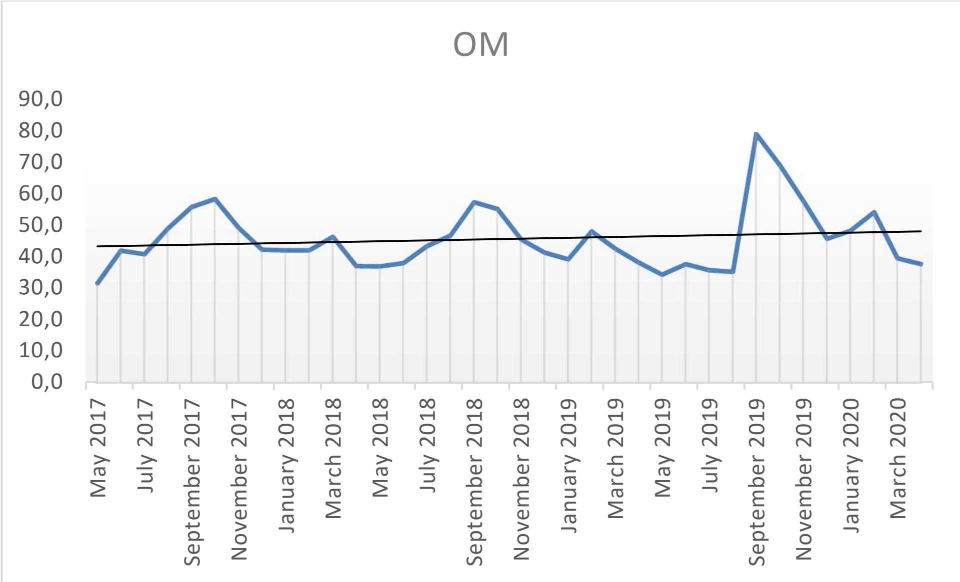


Figure 11: Water extents in km² for Oddar Meanchey over the period of three years with a linear trend.

3.1.4 Usage, Limitations and Constraints

All outputs of service 1 are based on Sentinel-1 and 2 imagery. For Sentinel-1 the revisit frequency is 12 days over Cambodia, means more speckle noise and in turn higher commission error rate in the monthly water masks especially at bare soils and very smooth surfaces.

As mentioned in 3.1.1, Sentinel-2 coverage is limited in 2017. During the wet seasons, high amounts of clouds can reduce the quality of the monthly composites. Sen2Cor misclassified clouds are also influencing the quality of the monthly image composites. Due to the shorter time frame of the wet season data gaps have a higher impact on the frequency-based classifications.

3.2 Service 2: Detailed Waterbody Inventory

3.2.1 Specifications

The technical specifications of the product comply with the properties issued in the technical proposal. The Small Water Features product is based on Very High-Resolution data, Pléiades 1-A and Worldview2. Each Province (Kouk mon and Bansay Reak) were delivered.

Table 6: Service 2 Product List

Product	**Count**	**Pixel size**	**Reference Period**	**File name**
Small water features mask	2	0.5m	2018-12-31 to 2019-01-26	KM_smallWaterFeaturesMask_2019.tif BR_smallWaterFeaturesMask_2019.tif

Table 7: Service 2 Product Specification

General	
Resolution and Data input	0.5m (Pléiades-1A, Worldview-2)
Geographic Projection	WGS84 / UTM zone 48N (EPSG : 32648)
Format	GeoTIFF
Datatype	Byte
Thematic information	
Classes and Codings	0 : no Small Water Features / 1 : Small Water Features / 255 : NoData
Accuracies	
Geometric positional accuracy	<1 meter
Overall thematic accuracy	85%
Geometric specifications	
Minimum Mapping Unit (MMU)	100m²
Minimum Mapping Width	1 meter

<u>Input Data</u>

List of input EO data used is described in Table 8.

Table 8: Description of the input data

Dataset	Spatial resolution	Spectral channels	Acquisition date	Spatial coverage
AOI 01 – Kouk Mon (270 km²)				
Pléiades 1-A	0.5m	4	2018-12-31	95%
Worldview-2	0.5m	3	2019-01-26	5%
Global Surface Water	30m	-	-	100%

AOI 02 – Bansay Reak (117 km²)				
Pléiades 1-A	0.5m	4	2019-01-19	75%
Pléiades 1-A	0.5m	4	2019-01-19	25%
Global Surface Water	30m	-	-	100%

Method

The methodology to extract small water features is based on Geographic Object-Based Image Analysis (GEOBIA).

The first task consists of collection and preparation of VHR data. Four scenes, Pléiades-1A and Worldview-2 were selected to cover the two area of interests. To reach the minimum horizontal resolution specification a pan-sharpening was carried out.

Secondly, a random sampling of ground truth based on Global Surface Water product and Open Street Map was carried out. However, representativeness of small water features and up-to-date sample of groundtruth was required to reach the product specification. Therefore, a manual fitting was performed to ensure a reliable sampling of ground truth. After a first classification iteration, a manual addition of counterexample was performed to improve the classification output.

The next step used a Machine Learning method to classify water elements based on the pansharpened multispectral scenes. This method extract texture features (Sobel, Pantex) and indices (NDVI) from EO data to allow water elements detection using a Random Forest classifier.

As expected with automatic detection of water using only VHR imagery, results showed some classification artefacts (see section 3.2.4). Therefore, a manual enhancement was carried out by focusing on specific omission or commissions artefacts like flooded fields, roads or built-up.

3.2.2 Quality Control and Validation

Classification correctness should be evaluated using misclassification rate and/or misclassification matrix. Thematic accuracy cannot be subjected to an exhaustive check. A thorough thematic assessment would imply a very time-consuming work. Misclassification rate is estimated by sampling and product information is compared to reference data.

The aim of the first paragraph is to provide a description of suggested procedures for a scientifically and statistically sound sampling scheme for assessing the thematic quality of the Water Presence products obtained for each area of interest.

Thus, thematic accuracy assessment has three components: (i) the sampling design, (ii) the response design and (iii) the estimation and analysis procedures.

The **stratification** and the sampling design primarily consist in selecting an appropriate sampling frame and sampling units. These sampling units can either be “defined on a cartographic representation of the surveyed territory” (Gallego, Area Frames for Land Cover Estimation: Improving the European LUCAS Survey, 2004[1]), in which case it is an area frame, or on a list of the features. According to this study, area frames give a better representation of the population as the spatial dimension is kept. In an area frame, sample units can be points, lines (often referred to as transects) or areas – often referred to as segments, described in (Gallego, Sampling

[1] Gallego, J., 2004. Area Frames for Land Cover Estimation: Improving the European LUCAS Survey., in: Proceedings of the 3rd World Conference on Agricultural and Environmental Statistical Application, Cancun, Mexico. pp. 2–4.

Frames of Square Segments, 1995[1]). For most cases, point samples will be used, but areas or segments may be used in specific cases such as when not only thematic accuracy needs to be reported, but also the geometry of mapped objects. Polygons have also the drawback of being specific to a single map. **Points are considered as the most appropriate unit for our products.**

Sampling design refers to the protocol whereby the samples are selected. A probability sampling design is preferred for its objectivity. "Simple random, stratified random, clustered random and systematics designs are all examples of probability sampling designs" (Stehman & Czaplewski, 1998[2]). Even though a simple random design is easy to implement, its main drawback lies in the fact that some portions of the population may not be adequately sampled. Cluster sampling is often used to reduce the costs of the collection of reference data but does not resolve geographic distribution problems. A systematic approach would solve this problem, yet it is not appropriate if the map contains cyclic patterns. A stratified approach consists in allocating a pre-defined number of samples per land-cover class. As explained in Stehman's paper, stratification ensures that each class is correctly represented.

The validation approach chosen combines **random and stratified approaches** and benefits from the advantages of both.

The stratification is applied based on a series of omission and commission strata:

- Commission: Small Water Features
- Omission: no Small Water Features

The **response design** "is the protocol for determining the reference land cover classification of a sampling unit" (Stehman & Czaplewski, 1998).

The datasets against which the interpretation is performed are divided in two main groups, guiding data and reference data. The guiding data used in the production of the classifications are VHR Pléiades 1A and Worldview-2 data. The guiding data are mandatory due the high seasonality of the areas of interest. The reference data provide more spatial details and stronger landscape context to the assessment. The available reference data are:

- Bing maps image / cartography layer
- Google Earth image / cartography data

The Table 9 gives a summary of the validation exercise and the accuracy assessment of the Water Presence products on each area of interest (Kouk Mon and Bansay Reak). It shows the Producer Accuracy and the User Accuracy related to the Omission and Commission Errors.

Table 9: Summary of the accuracy assessment for service 2

AOI	Producer Accuracy	User Accuracy
Kouk Mon	100.00%	95.92%
Bansay Reak	97.30%	100.00%

The accuracy assessment shows very satisfying results with very few omission and commission errors.

[1] Gallego, J., 1995. Sampling Frames of Square Segments (No. EUR 16317), Official Publication of the European Communities.
[2] Stehman, S. V., & Czaplewski, R. L. (1998). Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote sensing of environment*, *64*(3), 331-344.

3.2.3 Outputs

The outputs for Service 2 consist of Water Presence Products for each area of interest (Kouk Mon and Bansay Reak).

<u>Delivery package:</u>

- 2 water presence raster layers (one for each area of interest)
- INSPIRE compliant metadata .xml file

The Water Presence layers generated for both area of interest are presented in Figure 12 and Figure 13.

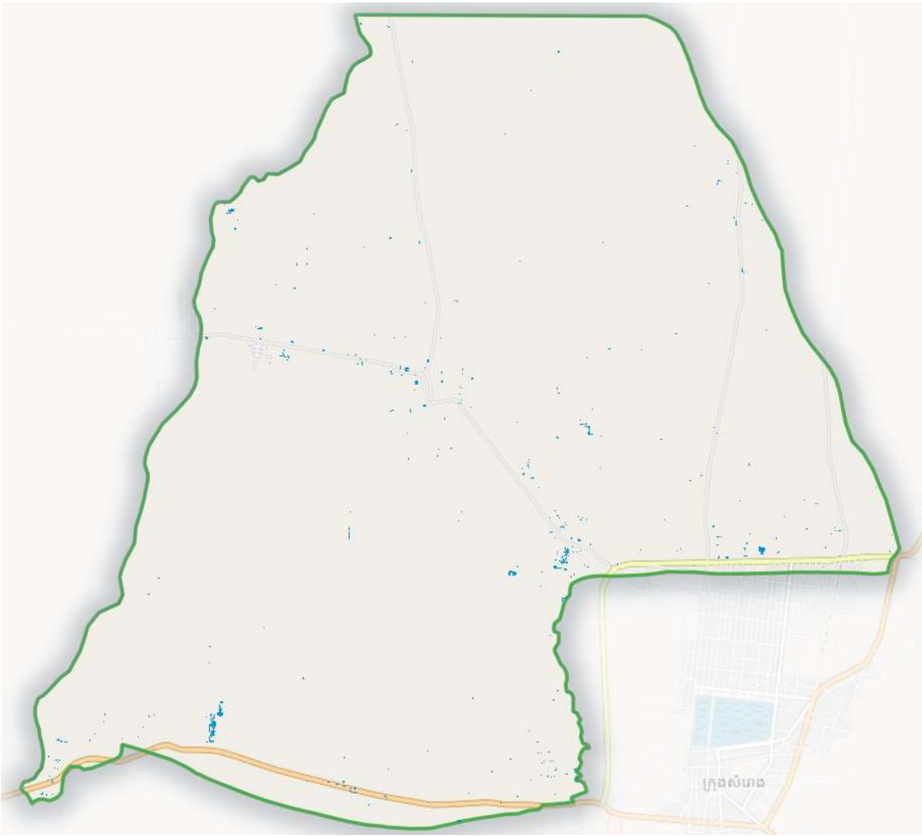


Figure 12: Very High-Resolution Water Mask over Bansay Reak area (blue) background : OSM

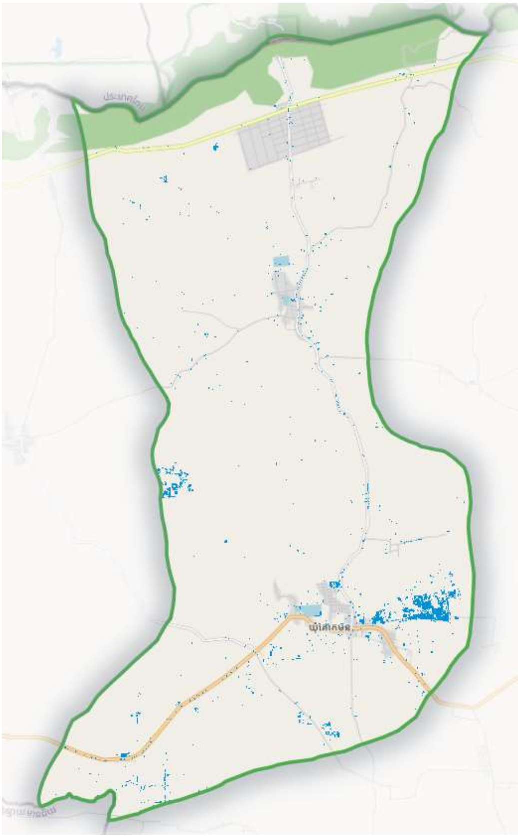


Figure 13: Very High-Resolution Water Mask over Kouk Mon area (blue) background : OSM

<u>Cost Benefit Analysis:</u>

Based on the two AoIs of Service 2, representing a total of 388km², we will try in this Cost Benefit Analysis to extrapolate the cost for a production at Service 1 area level and National level. First of all, the VHR EO data availability is the key to a large-scale production. In Service 2 AoIs demonstration case, we found multiple coverage of VHR imagery during the year, but unfortunately, only for dry season. The cloud coverage during the rainy season is too important to have a "usable" EO data to detect water features. Based on VHR imagery as soon as possible after the wet season (in our case December and January), we produce semi-automatically a water feature map. All the pre-processing (EO data preparation), classification (object-based approach) and postprocessing (filtering, formatting) are automatic steps. The ground truth collection, based on photointerpretation and usage of ancillary data, and manual enhancement are a manual part. In case of large-scale production, the ground truth data collection could be drastically reduced by mutualising the training for homogeneous area (same date of VHR coverage and same type of water features). The manual enhancement step is the only step that is linearly increasing with the size of the area to be produced. However, based on our experience on the service 2 area, we can conclude that even without manual enhancement, the results are pretty good. Omissions and commissions are due to lack of SWIR bands in VHR imagery, information useful to compute NDWI and to help water detection.

Based on Service 2 experience and the original financial proposition for EOC0010, we provide an estimation of the production of "Water Mask" based on VHR imagery for Service 1 area and at National scale (see Table 10).

Table 10: Cost Benefit Analysis: cost estimation for Service 1 area and National area.

Activities		Service 2 (388km²)		Service 1 (18.995km²)		National (181.035km²)	
		Quantity	Cost	Quantity	Cost	Quantity	Cost
EO data	VHR cost (unit km²)	400	3 200 €	19000	152 000 €	181000	1 448 000 €
Coordination		10	1 300 €	245	31 850 €	467	60 710 €
Development	Method	40	4 000 €	20	2 000 €	30	3 000 €
	Prototype	70	7 000 €	10	1 000 €	10	1 000 €
	Validation	35	1 750 €	10	500 €	10	500 €
Production	GT preparation	5	250 €	122,5	6 125 €	233,5	11 675 €
	EO data preprocessing	1	50 €	24,5	1 225 €	46,7	2 335 €
	Classification / postprocessing	2	100 €	49	2 450 €	93,4	4 670 €
	Manual enhancement	20	1 000 €	980	49 000 €	9340	467 000 €
	Validation	2	200 €	98	9 800 €	467	46 700 €
Total		185	18 850 €	1314	255 950 €	10230,6	2 045 590 €

3.2.4 Usage, Limitations and Constraints

The provided water features maps for the AoIs should allow users to have an exhaustive inventory of small water features (bigger than 100m²).

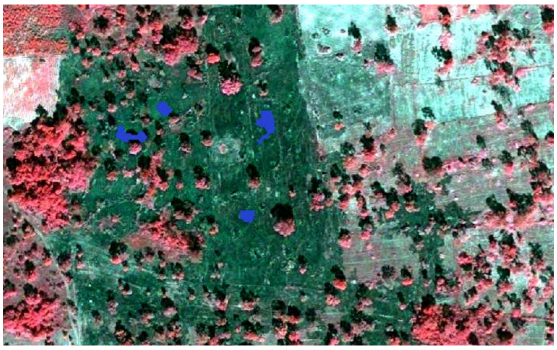
Due to the fact that this 2 AoIs are in a region mainly covered by clouds in wet season, it was decided with GIZ to produce the water feature maps using the closest data after the wet season, in our case, in December 2018 and January 2019. The water feature maps delivered is thus not representative to all the water features that could appear during the wet season, only a basis on water presence.

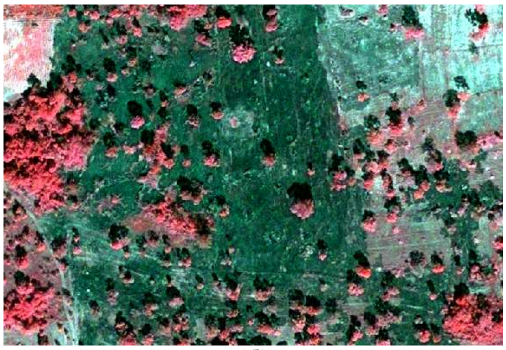
Due to the lack of some spectral information in VHR EO data (compared to Sentinel-2 for example), water detection is less perfect, and some confusion could occur, for example with building, inundated fields and roads (see **Error! Reference source not found.**). However, the final accuracy for both AoIs are significantly above the thematic accuracies requested.

Figure 14: Illustration of omissions and commissions before manual enhancement step

(a)

(b)


(c)


(d)

water mask in blue over VHR imagery (a) and (c) and respectively only the VHR image in (b) and (d).

APPENDIX A: BIBLIOGRAPHY

Gallego, J., 2004. Area Frames for Land Cover Estimation: Improving the European LUCAS Survey., in: Proceedings of the 3rd World Conference on Agricultural and Environmental Statistical Application, Cancun, Mexico. pp. 2–4.

Gallego, J., 1995. Sampling Frames of Square Segments (No. EUR 16317), Official Publication of the European Communities.

GIZ (2019): Transboundary Water Resource Management in the Lower Mekong Basin – Joint Project Flood and Drought Thailand and Cambodia Sub Basins 9T/9C

Ludwig et al. (2019). A highly automated algorithm for wetland detection using multi-temporal optical satellite data. Remote Sensing of Environment, 224, 333-351

Martinis et al. (2009). Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data. Natural Hazards and Earth System Sciences, 9(2), 303-314

Merciol et al. (2019). GEOBIA at the Terapixel Scale: Toward Efficient Mapping of Small Woody Features from Heterogeneous VHR Scenes. ISPRS Int. J. Geo-Inf. ***2019***, 8*, 46*.

Ramsar Convention, 2016.An Introduction to the Ramsar Convention on Wetlands, 7th ed. Ramsar Convention Secretariat, Gland, Switzerland.

Rennó, et al. (2008). HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. Remote Sensing of Environment, 112(9), 3469-3481

Roberts et al. (2017). High-Dimensional Pixel Composites from Earth Observation Time Series. IEEE Transactions on Geoscience and Remote Sensing, 6254-6264

B. Sanders (1999) A very interesting technical report, Version 2.1 (Ref. C2-TCN-A-01-AA), Institute Name.

A. Smith and B. Anderson (2012): The Advanced Module of the Advanced System, presented at A Very Important Conference, City Name, Country, 6–10 September.

A. Smith, B. Anderson, C. Johnson and D. Jackson (2012): An evaluation of some work about some other work, Very Important Journal, 13(2), pp. 115–124.

Stehman, S. V., & Czaplewski, R. L. (1998). Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote sensing of environment*, *64*(3), 331-344

APPENDIX B: ANALYSES - WATER AREA VALUES BY MONTH

Water extents in km² by months for both provinces for the entire 3-year-series.

Kampong Thom		
Date	*Pixels*	*Area (km2)*
May 2017	4935714	**493.6**
June 2017	5794544	**579.5**
July 2017	7558079	**755.8**
August 2017	11239691	**1124.0**
September 2017	11952568	**1195.3**
October 2017	11871853	**1187.2**
November 2017	11192768	**1119.3**
December 2017	9490275	**949.0**
January 2018	7132465	**713.2**
February 2018	5664745	**566.5**
March 2018	5101400	**510.1**
April 2018	4673195	**467.3**
May 2018	4696757	**469.7**
June 2018	5536156	**553.6**
July 2018	7528438	**752.8**
August 2018	12442056	**1244.2**
September 2018	15116963	**1511.7**
October 2018	14521850	**1452.2**
November 2018	12862870	**1286.3**
December 2018	8851680	**885.2**
January 2019	6120121	**612.0**
February 2019	5072746	**507.3**
March 2019	4663472	**466.3**
April 2019	4389472	**438.9**
May 2019	4188638	**418.9**
June 2019	4318005	**431.8**
July 2019	4187694	**418.8**
August 2019	6295395	**629.5**
September 2019	12072662	**1207.3**
October 2019	12830763	**1283.1**
November 2019	9018986	**901.9**
December 2019	6120993	**612.1**
January 2020	4931097	**493.1**
February 2020	4455094	**445.5**
March 2020	4038272	**403.8**
April 2020	3673014	**367.3**
May 2020	3477198	**347.7**
June 2020	3627490	**362.7**
July 2020	3625140	**362.5**

Oddar Meanchey		
Date	*Pixels*	*Area (km2)*
May 2017	315514	**31.6**
June 2017	418978	**41.9**
July 2017	408091	**40.8**
August 2017	489490	**48.9**
September 2017	557403	**55.7**
October 2017	583200	**58.3**
November 2017	491846	**49.2**
December 2017	422674	**42.3**
January 2018	419874	**42.0**
February 2018	420339	**42.0**
March 2018	462541	**46.3**
April 2018	370051	**37.0**
May 2018	369688	**37.0**
June 2018	379803	**38.0**
July 2018	433850	**43.4**
August 2018	466749	**46.7**
September 2018	573149	**57.3**
October 2018	551627	**55.2**
November 2018	454511	**45.5**
December 2018	412883	**41.3**
January 2019	391726	**39.2**
February 2019	480818	**48.1**
March 2019	425670	**42.6**
April 2019	381115	**38.1**
May 2019	343108	**34.3**
June 2019	376143	**37.6**
July 2019	357124	**35.7**
August 2019	351586	**35.2**
September 2019	789358	**78.9**
October 2019	689822	**69.0**
November 2019	577547	**57.8**
December 2019	456664	**45.7**
January 2020	482873	**48.3**
February 2020	541753	**54.2**
March 2020	394958	**39.5**
April 2020	376106	**37.6**
May 2020	415895	**41.6**
June 2020	438856	**43.9**
July 2020	467362	**46.7**